DOCUMENT RESUME

ED 077 932 TM 002 727

AUTHOR	Haynes, Judy L.; Impara, James C.
TITLE	Florida Statewide Assessment Program 1971-72 Technical Report; Section 1: Introduction, Procedures, and Program Recommendations.
INSTITUTION	Florida State Dept. of Education, Tallahassee. Bureau of Planning and Evaluation.
PUB DATE	72
NOTE	23p.
AVAILABLE FROM	Evaluation Section, Department of Education, Tallahassee, Florida 32304
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Educational Accountability; *Evaluation Methods; Grade 2; Grade 4; *Program Descriptions; *Reading Achievement; *Reading Skills; Reading Tests; State Programs; *Student Evaluation; Technical Reports
IDENTIFIERS	*Florida

ABSTRACT

The first section of a four-part technical report of Florida's statewide program for assessing reading-related skills in grades 2 and 4 provides an introduction to the program, a description of procedures used, and recommendations regarding program operation. Program background, design, and responsibility for assessment activities are discussed in the introduction. The procedures discussed are: selection of statewide objectives, pupils and pupil characteristics, selection of random sample, assessment instruments, testing conditions, distribution of materials, completion of local testing, and test scoring. The recommendations on program operation involve special target groups, representativeness of sampling, test materials, instructions and general information, and packaging and distribution of test materials. Florida's Educational Accountability Act of 1971, the legislative impetus for this assessment program, is provided in an appendix. (For related documents see TM 002 724-726.) (KM)

FILMED FROM BEST AVAILABLE COPY

ED 077932
TM 002 727

10.19

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Florida Statewide Assessment Program

TECHNICAL REPORT

Section 1

Introduction, Procedures, and Program Recommendations

1971-1972

DEPARTMENT OF EDUCATION
TALLAHASSEE, FLORIDA
FLOYD T. CHRISTIAN, Commissioner

Section 1, "Introduction, Procedures, and Program Recommendations," of the Technical Report was prepared by Judy L. Haynes and Dr. James C. Impara of the Evaluation Section, Bureau of Planning and Evaluation, Division of Elementary and Secondary Education, Department of Education. Requests for information or copies should be sent to the Evaluation Section, Department of Education, Tallahassee, Florida 32304.

This public document was promulgated at an annual cost of $1,906.21 or $.76 per copy to inform interested parties of the 1972 Florida Statewide Assessment results and to provide information with which these results can be interpreted pursuant to Chapter 229.57(3)(c) Florida Statutes.

ED 077932

Florida

Statewide Assessment Program

1971-72 TECHNICAL REPORT

Section 1

INTRODUCTION, PROCEDURES, AND PROGRAM RECOMMENDATIONS

Florida Department of Education

Division of Elementary and Secondary Education

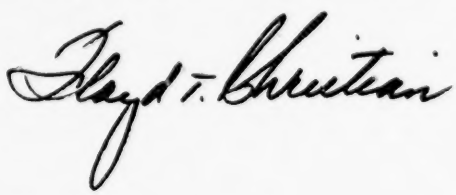
Bureau of Planning and Evaluation

FOREWORD

The Florida Department of Education is currently implementing a statewide assessment program which will provide in-depth information about the strengths and weaknesses of Florida's educational system. One of the most important features of this assessment program is its measurement of specific educational objectives which identify the skills Florida students should achieve from their educational experience.

The information about student achievement will enable educators to pinpoint weak spots and redistribute educational resources to achieve better results. The ultimate effect of this effort will be improvement in the degree to which our school system prepares students to function in society.

Developing an assessment program of this scope has been a monumental effort for the Department of Education. I am proud that the Department can present the procedures and results of this year's assessment program as a step toward improving the state educational system.



Floyd T. Christian
Commissioner

PREFACE

One provision of the Educational Accountability Act of 1971 was the establishment of a Statewide Assessment Program which would measure the degree of student achievement of statewide educational objectives. As the first step in implementing the assessment program, in 1971-72 a sample of second and fourth graders in each school in the State were tested on their achievement of selected reading-related skills. Approximately 53,000 students, or twenty percent of the students in each grade, were tested on the statewide objectives.

These objectives, chosen by teachers and other educators throughout the State, identified a number of important reading-related skills. Achievement of the objectives was measured through objective-referenced tests; that is, each objective was measured by one or more items.

This, the first public report of the Statewide Assessment Program, outlines the background of Florida's Assessment and Accountability Programs. In addition, it describes the procedures, results and recommendations of the 1971-72 assessment of selected reading-related skills.

The report has been prepared in two forms: a brief Capsule Report which summarizes the key results and recommendations of the State's performance on the assessment; and a multi-section Technical Report which describes the entire assessment program. The sections of the Technical Report are:

Section 1: Introduction, Procedures, and Program Recommendations

Section 2: Statistical Information

Section 3: Statewide Results and Recommendations

Section 4: District Interpretations

You are reading Section 1 of the Technical Report. The other sections of the Technical Report and the Capsule Report are available upon request from the Evaluation Section, Department of Education, Tallahassee, Florida 32304. When requesting the Technical Report, please indicate which sections you wish to receive. For Section 4, "District Interpretations," the subreport for each district is bound separately, so please identify which district's interpretation you would like.

TABLE OF CONTENTS

Chapter

I. INTRODUCTION 1

Background of Assessment Program 1
Design of 1971-72 Assessment 3
Responsibility for Assessment Activities 4

II. PROCEDURES 5

Selection of Statewide Objectives 5
Pupils and Pupil Characteristics 6
Selection of Random Sample 7
Assessment Instruments 7
Testing Conditions 10
Distribution of Materials 10
Completion of Local Testing 11
Test Scoring 11

III. RECOMMENDATIONS ON PROGRAM OPERATION 13

Special Target Groups 13
Representativeness of Sampling 14
Test Materials 15
Instructions and General Information 15
Packaging and Distribution of Test Materials 16

Appendix

A. THE EDUCATIONAL ACCOUNTABILITY ACT OF 1971 17

CHAPTER I

INTRODUCTION

In 1971, the Florida Legislature passed the Educational Accountability Act of 1971, which required the establishment of an Accountability Program for the State's educational system. As Commissioner of Education Floyd T. Christian explained, there are three components to the Accountability Program: "1) the evaluation of pupil performance and behavior, 2) the identification of factors in the teaching-learning situation which have a direct bearing on performance and behavior, and 3) an attempt to relate the cost of these factors to learning outcomes."[1]

The Florida Accountability Program is a comprehensive system for evaluating the State's educational system. Its function is to gather information about the three components--pupil performance, educational procedures, and costs--and disseminate it to educational managers in a usable, timely manner. The Statewide Assessment Program was established to handle the evaluation of pupil performance and behavior.

The Educational Accountability Act required that reading be the first subject area assessed. Reading was selected because it is, as Commissioner Christian said, "the first and most basic of subjects," and is the skill required as a prerequisite for the vast majority of educational activities. Since a primary purpose of the first statewide assessment was to establish and evaluate assessment procedures, testing in 1971-72 was limited to second and fourth grade reading-related skills, with plans to expand assessment to other grade levels and subject areas in subsequent years.

BACKGROUND OF ASSESSMENT PROGRAM

According to Commissioner Christian,[2] the move toward educational accountability began in 1968, when the Legislature enacted a law requiring the Department of Education to expand its capabilities to effect constructive educational change, improve the quality of education, and make the State's educational

---

[1]Floyd T. Christian, "A Message to Florida Teachers," Florida Education Tabloid, March, 1972, p. 10.

[2]Ibid.

program more effective and economical. Legislation passed in 1970 (Chapter 70-399, Laws of Florida) required the Commissioner of Education to develop evaluation procedures to assess objectively the progress of students at various grade levels and in the various educational programs of the public schools. The plans for implementing this legislation are outlined in the Plan for Educational Assessment in Florida, which describes the rationale and plans for the first year of operation of the Florida Accountability Program.[1]

The 1971 Legislature adopted the plans in the Educational Accountability Act of 1971 (Chapter 229.57, Florida Statutes).[2] Major provisions of the Act relate to:

1. Establishment of "basic, specific, uniform statewide educational objectives for each grade level and subject area including, but not limited to, reading, writing and mathematics, in the public schools."

2. Development of a uniform, statewide system of assessment to determine periodically pupil status, pupil progress, and the degree of achievement of established educational objectives.

3. Development of procedures capable of providing information for 1) analysis of the costs associated with public education programs and 2) analysis of the differential effectiveness of instructional programs.

4. Annual public reporting of assessment results by grade and subject area for each school district and for the State.

The Act further indicated that these provisions would apply to the subject area of reading for the 1971-72 school year; the subject areas of reading, writing, and mathematics for the 1972-73 school year; and additional subject areas for 1973-74.[3]

The Statewide Assessment Program is responsible for carrying out provisions 2 and 4, the development of a statewide system of assessment and the annual reporting of assessment results. Other programs have been established to accomplish the remaining provisions of the Accountability Act (see p. 4).

---

[1]Copies of the Plan for Educational Assessment in Florida are available from the Evaluation Section, Department of Education, 117 Miles-Johnson, Suite 10, Tallahassee, Florida 32304.

[2]Appendix A contains the complete Educational Accountability Act of 1971.

[3]Because of the complexity of developing workable procedures, information on costs and instructional programs was not collected during 1971-72. These components of the Accountability Program will become operative at a later date.

DESIGN OF 1971-72 ASSESSMENT

A critical factor in the design of the 1971-72 assessment program was the decision that the primary goal for this year would be the development and evaluation of assessment materials and procedures. Florida had some experience in statewide testing programs in the ninth and twelfth grades, using traditional types of tests. However, the State had no experience in measuring specific educational objectives in a variety of grades and subject areas through a statewide program. Consequently, the assessment plans for 1971-72 called for the use of objective-referenced tests designed specifically to measure the degree of achievement of Florida's statewide educational objectives.

A second factor in the design of the program was the decision to test a sample of students, rather than every student. The Educational Accountability Act required that results be reported for every district and for the State. Sampling was chosen because it can provide district and State level information reliably and much more efficiently and economically than testing every student. This factor was particularly important in 1971-72, when the budget was small and many of the procedures had not been tried out previously.

The particular sampling technique used (multiple-matrix sampling) maximized the number of objectives which could be tested, while minimizing the number of students and length of testing time involved. Multiple-matrix sampling means that only a portion of the students in each school are tested, and each student who is tested takes only a portion of the items which constitute the complete test. This means that each student spends less time in the testing situation and fewer students are tested, while the data allows us to make reliable statements about student performance for the district and the State. (A complete discussion of multiple-matrix sampling is contained in Appendix A of the Plan for Educational Assessment in Florida.)

Another important decision related to the number of students to be tested. In order to gain experience with assessment at different grade levels, assessment was originally planned to cover reading skills in grades two, four, seven, and ten by selecting randomly approximately 10,000 students for testing throughout the State. It was later decided that the assessment data should also supply data for the State's school accreditation program. In order for the data for accreditation to be reliable, a minimum of 3-4 students had to be tested on each objective in each school. This drastically increased the number of students to be tested, so because of budget limitations (and delays in development of assessment material), grades seven and ten were not assessed during 1971-72. Nevertheless, almost 53,000 students were assessed, as opposed to the initial estimate of 10,000.

The assessment plan which resulted from these decisions required the selection of specific educational objectives for reading in grades two and four, the development of exercises to measure each of the selected objectives, and the testing of a random sample of the State's second and fourth graders. Throughout the year, efforts were made to evaluate assessment procedures and determine what, if any, alternative approaches should be tried in 1972-73 assessment.

RESPONSIBILITY FOR ASSESSMENT ACTIVITIES

Responsibility for administering the Statewide Assessment Program is assigned to the Evaluation Section in the Division of Elementary and Secondary Education, Department of Education. The Evaluation Section is essentially a research and testing unit, so some activities--such as the development of educational objectives and the interpretation of assessment results in specific subject areas--are the responsibility of other sections within the Department.

The network set up for the administration of the 1972 reading assessment included the Evaluation Section, a district coordinator for accountability in each of Florida's 67 school districts, and a school coordinator and test administrators in each school where pupils were tested. The assessment program depends heavily on the involvement of the district and school personnel, both to operate the program successfully and to make it responsive to the needs of individual districts and schools.

The district coordinators were selected by their respective superintendents and were responsible for conducting assessment activities within their district. They handled the dissemination of information and materials to the school coordinators and test administrators, monitored all assessment-related activities, and returned all assessment materials to the Department of Education.

One person in each school, the school coordinator, was assigned to supervise activities within the school. The school coordinator worked with the district coordinator, trained the test administrators, handled all preparations at the school, supervised the actual testing, and prepared the materials for return to the district coordinator.

Each school had one or more test administrators. The test administrator read the test instructions and questions to the students, collected the test books after the students were finished, and filled in the pupil data requested on the back cover of the test.

Because many of the responsibilities for conducting assessment were placed on the local district and school level, the district coordinators were frequently asked for recommendations related to the program. At the first coordinators' conference, held in September, 1971, they provided recommendations relating to division of responsibility between district coordinators, school coordinators, and test administrators; procedures for student sampling; and details of test administration. A second meeting was held in March, 1972, to critique the assessment program. Input was also obtained through the use of questionnaires, and, in some cases, phone calls to representative districts.

CHAPTER II

PROCEDURES

The Educational Accountability Act of 1971 specified that the assessment program should be objective-referenced; that is, testing should be for the purpose of relating student achievement to established statewide educational objectives. Florida did not have previous experience with a statewide objective-referenced testing program, which made it necessary to develop a complete set of test exercises and procedures in less than twelve months.

In anticipation of the Educational Accountability Act of 1971, assessment-related activities began with the development of exercises to assess reading in the 1971-72 school year. On March 1, 1971, the Research and Development Section of the Department of Education contracted with the Center for the Study of Evaluation (CSE) at the University of California in Los Angeles to provide Florida with instruments to measure selected reading objectives in grades two, four, seven, and ten.

Assessment for grades two and four was originally scheduled for November, 1971, and for grades seven and ten in spring, 1972. However, delays in the completion of assessment exercises forced postponement of testing in grades two and four until February, 1972. Assessment in grades seven and ten was not conducted during 1971-72, for the reasons discussed on page 3.

SELECTION OF STATEWIDE OBJECTIVES

CSE provided a catalogue of reading objectives, from which Florida reading specialists and classroom teachers selected the objectives to be measured in the 1971-72 assessment program. The individuals who participated in the selection process, 112 reading specialists and 236 classroom teachers from throughout Florida, were chosen by their district superintendents.

The participants were first asked to establish the level at which each objective should be mastered (levels were defined as grades two, four, seven and ten). After objectives had been assigned to the appropriate level, the participants were asked to select those objectives which were of highest priority. "Highest priority" was operationally defined as objectives which virtually all students should attain at the level specified.

The survey provided reasonable estimates of appropriate levels for various objectives, but few objectives were rejected in the priority listing, since most of the 1200 objectives in the catalogue were determined to be of highest priority. Consequently, the Department of Education utilized an outside consultant in reading to assist in reducing the list to a more manageable size.[1] The consultant used the results of the selection process to eliminate those objectives which were selected less frequently.

The reduced lists of objectives for grades two and four were presented to, and adopted by, the Florida State Board of Education on September 14, 1971, as statewide educational objectives. They were labeled high priority objectives for reading in Florida for students seven and nine years of age,[2] to establish a precedent for eventually shifting away from the traditional grade structure. For 1971-72, age seven was defined operationally as grade two and age nine was defined as grade four, in order to simplify the process for selecting the sample. In future years, students will be selected by a combination of grade and age.

As part of the overall assessment and accountability strategy, the Department of Education is in the process of preparing catalogues of objectives in a variety of subject matters and grade levels. In order to conduct the 1971-72 assessment, the Department obtained objectives for grades two and four from CSE. Subsequently, the Department contracted with Florida State University to provide a complete catalogue of reading objectives which would cover an increased number of reading-related skills and expand the objectives to include grades K-12. New statewide objectives will be selected from the revised catalogue.

PUPILS AND PUPIL CHARACTERISTICS

For the 1971-72 reading assessment, second graders were defined as (1) if graded, those students in the second grade, or (2) if ungraded, those students in their second year of school without kindergarten or their third year of school with kindergarten. Fourth graders were those students in the fourth grade, if graded; or if ungraded, those students in their fourth year of school without kindergarten, or fifth year of school with kindergarten. During 1971-72, there were 118,858 second graders and 121,843 fourth graders in Florida schools. Each school which had 12 or more students in grades two and/or four was included in assessment.

---

[1] The Department of Education did not have a reading consultant at that time.

[2] See "1971-72 High Priority Objectives for Reading in Florida, Ages 7 and 9," which is included as Appendix A of Section 2. The document is also available from the Evaluation Section, Department of Education, 117 Miles-Johnson Building, Suite 10, Tallahassee, Florida 32304.

The testing population included most students in the second or fourth grade, including Educable Mentally Retarded; children whose predominant language was not English; children with special physical or emotional problems; children who were repeating a grade; and, for the fourth grade test, those who had skipped grade three. The only group of students who were excluded from the testing population were those students classified as Trainable Mentally Retarded.

SELECTION OF RANDOM SAMPLE

The sample of students to be tested was selected by the school principal, coordinator of assessment, or other designated official at least one week prior to testing. In some districts, selection of samples was done under the supervision of the district coordinator. A detailed set of instructions for selecting the sample was sent to each school.[1] The instructions identified the steps to be followed in selecting the sample and contained two tables: one for determining the number of students to be tested, based on school size; and the other a table of random numbers for selecting the students to be tested.

The first step in the selection process was the preparation of an alphabetical master list of all students in the school who were in the grade being assessed. Every name on the list was assigned a number. The person selecting the sample referred to the first table to determine how many students would be tested in his school, and then went to the table of random numbers and selected that many numbers from the table. Each number selected matched a number on the master alphabetical list of students, and thus identified the students to be tested. The first number (student) selected was assigned to take Form A, the second to Form B, the third to Form C, the fourth to Form D, the fifth to Form A, and so on, until the appropriate number of students had been assigned to forms.

The instructions indicated that if any of the students selected in the original random sample were absent on the day of testing, a replacement should be selected by taking the next student on the master list who was not already chosen for testing.

ASSESSMENT INSTRUMENTS

With multiple-matrix sampling, each student takes only a portion of the items, and scores are combined to indicate how all students would have performed if they had taken all of the items. For the 1971-72 reading assessment, four different test forms were used for assessment in each grade. Each form measured approximately one-fourth of the selected statewide objectives.

---

[1]Copies of the Sampling Procedure used for the 1971-72 assessment of reading-related skills can be obtained from the Evaluation Section, Department of Education, 117 Miles-Johnson, Suite 10, Tallahassee, Florida 32304.

Test Development

CSE contracted to supply test items to measure each of the selected statewide educational objectives; arrange the test items into four forms for each grade; prepare the administration directions; and deliver all materials in camera-ready format. Initial delivery of assessment items for grades two and four was made on September 1, 1971, with the understanding that State Board of Education action on the proposed statewide objectives might require some changes in the test contents. The State Board adopted the objectives as proposed, so no changes were required for this reason.

However, because of certain technical difficulties, modification and replacement of test items continued until November 12, 1971. In addition, there were delays in receiving answer keys, objective-item correspondence keys, and general instructions for administration. In order to meet the February testing schedule, the tests and instruction manuals were sent to the printer immediately after receipt of the last items from CSE.

Test Contents

The reading-related skills measured on the tests were listening comprehension, auditory and visual perception and discrimination, identification of phoneme-grapheme correspondence, syntactical structure, word processing, meaning, study skills, and reading comprehension. A description of each skill and an example of the type of item used to measure it is given in Chapter 1 of Section 3 of the Technical Report.

The item pool consisted of 116 items for second grade objectives and 291 items for fourth grade objectives. The number of items per grade exceeded the number of objectives because some objectives were measured by two or more items. No items were submitted for twelve second-grade objectives and twenty-seven fourth grade objectives, so these skills were not tested in the 1971-72 assessment. The objectives for which there were no items are identified in Appendix A of Section 2.

Test Validation

A study of the content validity of the reading assessment items was conducted on November 29-30, and December 1, 1971. Since the tests were being printed at that time, the purpose of the evaluation was to determine the use to be made of each item in reporting the results. The objectives of the review board were:

(1) to evaluate each item to determine if it measured the stated objective;

(2) to evaluate each item to determine if it had technical errors;

(3) based on the above evaluation, to decide if student performance on the item should be tabulated and included in the report of results;

(4) to decide, in cases where two or more items measured an objective, how many items should be answered correctly to indicate achievement of the objective.

Based on the group's recommendations, it was decided that performance on items which lacked content validity would not be reported. Performance on items with technical difficulties would not be reported, if the judgment of the panel were subsequently supported by an analysis of student performance on the item. The group recommended that results on approximately 32% of the objectives should not be reported, because the items did not measure the objectives or had serious technical difficulties.[1] In addition, they noted that reading-related skills such as intelligence and listening were measured to a greater extent than specific reading skills, and recommended that the title of the test be "A Preliminary Test of Selected Reading-Related Skills."

Participants in the validation study included Dr. Charles Mangrum, University of Miami; Dr. Edward Fry, Rutgers University; Dr. Edwin Smith, Dr. Lawrence Hafner, and Dr. Howard Stoker, Florida State University; Dr. Tom Fillmer, University of Florida; Dr. Voncile Mallory, Florida Atlantic University; Dr. Elsie Wallace, Florida A & M University; Mrs. Lorene Lawson, Pinellas County; Mr. Richard White, Dade County; and Mrs. Charlotte White, Mrs. Kittie Mae Taylor, Mrs. Blanche McMullen, Dr. James Impara, and Dr. Martha Cheek, Florida Department of Education.

Test Format

The tests developed by CSE were not designed to permit machine scoring. Instead, students were required to mark in the answer booklet. The primary response mode was placing an "X" in a box either under or next to the correct response, in a multiple-choice format. However, some items required other responses, such as circling or underlining, or required the student to write words or numbers (see Table 1).

Table 1. Number of items with different types of responses on each grade test

	TYPE OF RESPONSES					
Grade	Marking An 'X'	Writing Words	Writing Numbers	Circling	Drawing Slashes	Underlining
Two	108	2	2	---	---	2
Four	254	15	8	9	4	1

[1] See Appendix A in Section 2 for a breakdown of why objectives were omitted from the reporting of results.

Answer keys for all items were developed by CSE. No partial scores were allowed; each item was scored as either right or wrong.

The tests for each grade consisted of four test forms--A, B, C, and D--with a different set of administration instructions for each form. For ease in identification of test forms, the cover of each form was a different color (four colors in all), and the administration instructions for each form were printed on paper of a color corresponding to the test booklet cover. Each booklet was assigned a unique five digit number for identification purposes, and this number (not the student's name) was used to identify scores. On the back of each test booklet were spaces asking for the student's name, age, sex and race/national origin, so that the sampling procedure could be verified. The test administrator was asked to fill in this information, and also to write in the school name and number, the time and date of testing, and his own name.

TESTING CONDITIONS

Each form of the test was given to a different group of children. The second grade forms required about an hour to complete, while the fourth grade forms, each consisting of two parts, took about two hours. Because of the length of the fourth grade tests, a 30-minute break was scheduled between the administration of parts I and II of the test. There was no break during the administration of the second grade test.

In order to accommodate schools on double sessions, tests were administered during the "first part of the normal school day." Because many instructions and questions had to be read aloud by the administrator, the forms could not be administered simultaneously. The second grade tests were given on Tuesday and Wednesday, February 8 and 9, 1972, and the fourth grade tests were given one week later, on February 15 and 16. Ninety-five percent of the schools tested on the four scheduled dates; the remaining schools tested a day later because of local conditions such as holidays and field trips. In both grades forms A and B were given the first day of testing, forms C and D the second day.

The tests were not timed, and the test administrator was instructed to allow enough time between questions so that all children could answer the question. The instruction manual indicated that thirty seconds of answer time should be adequate for most questions but that some which required written responses might take up to three minutes. As soon as all students had completed the question (or the time had elapsed), the administrator proceeded to the next question.

DISTRIBUTION OF MATERIALS

Packaging and mailing of assessment materials were accomplished through the efforts of the Evaluation Section and several employees hired on a temporary basis to assist with the packaging. Each package contained a predetermined number of copies of each test form, directions for the administration of each form, and instructions for selecting the random sample of students to be tested. The packages were mailed to the appropriate district coordinators, who distributed them to the schools.

After testing in both grades was completed, all test materials were returned to the district coordinator, who repacked and returned them to the Evaluation Section for scoring and analysis.

COMPLETION OF LOCAL TESTING

Included in the administration manual were two pages for test administrator and/or school coordinator comments. One page was for suggestions related to the procedures used in the 1971-72 statewide assessment. On the second sheet, they were instructed to explain any variances from the procedures outlined in the administration manual and the sampling procedure. The school coordinators were instructed to return all assessment materials, including administration manuals, sampling plans, master alphabetical lists, and unused tests.

TEST SCORING

The contract for test scoring was awarded to Software Programming and Associates, from Titusville, Florida. They examined all testing materials to discover if any test booklets or other materials were missing. Follow-up procedures were conducted to facilitate the return of missing materials.

Since the tests could not be machine-scored, the contractor arranged for hand-scoring of the tests, using answer keys provided by CSE. Scoring of tests began about March 2, 1972. When all tests were scored, the Department of Education conducted a study to determine the accuracy of the contractors' scoring. Overall, their error rate was less than one percent, but it was as high as 20% on one particular item. Items with excessive scoring errors were not reported in the results. The re-scoring study is described in Section 2, with further information on which items and objectives were omitted because of excessive scoring error.

CHAPTER III

RECOMMENDATIONS ON PROGRAM OPERATION

In commenting on the magnitude of the developmental effort required to launch the statewide assessment program, Commissioner Christian noted that "every element in the entire program had to be developed from 'scratch'. None of the materials, the procedures, or the processes used were available from other sources." Because the entire program had to be developed for the 1971-72 assessment in reading, each activity was examined carefully to identify those areas which needed revision or modification in order for assessment to be more successful in future years.

Suggestions for revision came primarily from four sources. (1) During testing, Department of Education staff visited schools in various parts of the State to observe the testing procedure first-hand and note difficulties. (2) Each test administration manual included a "Suggestion" sheet, and the responses of test administrators on this suggestion sheet were tabulated. (3) The district coordinators met in March, 1972, for the primary purpose of critiquing the assessment procedures and recommending improvements. (4) In addition, the Evaluation Section examined its own activities in such areas as contract monitoring, data analysis, and report preparation; and recommended changes which would make the operation of assessment more efficient.

This report describes some of the problems encountered during the first year of operation of the Florida Statewide Assessment Program and discusses the recommendations for eliminating or minimizing the problems. As would be expected, the same suggestions often came from all four sources, so the problems will be discussed by area, rather than by source.

SPECIAL TARGET GROUPS

There has been general concern throughout the State's educational community that the statewide assessment exercises did not accurately assess the achievement of special target groups of students (such as the Educable Mentally Retarded, children whose native language is not English, and children with special physical or emotional problems). It has been suggested that the test, the administration procedures, the performance criterion, and/or the reporting format should be adapted for these special groups.

In response to these suggestions, the Department of Education has been holding intradepartmental meetings to discuss the problems of assessment for special groups of children. A feasibility study will be conducted to determine what changes can be made to adapt assessment for the State's EMR students. It is anticipated that, if personnel and resources are available, these adaptations will be field-tested in 1972-73. Adaptations for other special target groups may be made in future years, based on a careful analysis of 1972-73 assessment procedures and results.

REPRESENTATIVENESS OF SAMPLING

In general, the use of multiple-matrix sampling was an effective and acceptable technique with which to conduct statewide assessment. Because of the number of objectives and items, a second grader who completed all items could have been subjected to as much as four hours of testing, while a fourth grader could have experienced up to eight hours of testing. Consequently, the use of multiple-matrix sampling significantly reduced the testing time for individual students and was generally less disruptive of the classroom than lengthier testing would have been.

The primary concern about the sampling procedure was not item sampling, but the concept of sampling students rather than testing everyone. This concern was especially apparent at the school level, where administrators felt that the random sample of students did not represent their school's population. Since by design the sampling plan was intended to be representative only at the district and State levels, it is possible that some school samples were not representative. However, as long as the data are not reported by school, and the sample is representative on a district basis, there are no problems with the use of student sampling. Preliminary indications are that the samples on a district and State level were representative. Special studies conducted voluntarily in two counties, Hillsborough and Escambia, show that the random sample was almost proportionate to the population of their districts, and it appears the sample was also representative for the State.

The success of this year's use of multiple-matrix sampling justifies the continued use of the technique, at least for the 1972-73 school year when the results will again be reported only on a district and State level. Although at the present time results are not reported by school, the need for such reporting may eventually arise. In order to report by school, the random sample would have to be representative on a school basis. The Department of Education is currently investigating alternative procedures to insure representativeness on a school level. Possible choices include changes in the basic sampling plan and a challenge system which would allow a school to be re-tested if it could demonstrate that the sample was not representative. The Department of Education is preparing a long-range assessment plan which will utilize random sampling in some grades and subjects and census testing in others.

TEST MATERIALS

Concerns about the test were basically of two types. The first was that some items seemed too advanced for the grade being assessed. Since the test is objective-referenced, the charge that some items were too difficult suggests that there is disagreement with some of the objectives measured by the test. The recommendations of the committee who examined the validity of the items also indicated certain weaknesses in the objectives. The problems encountered in the preparation of the 1971-72 statewide objectives were used as guidelines in an extensive revision of the catalog of reading objectives, which has just been completed. New statewide reading objectives have been selected and only a small number of the 1971-72 reading objectives will again be used for statewide assessment.

The second type of criticism related to the general preparation of the test. According to the test administrators, many of the drawings and illustrations were poor. Some of the directions were ambiguous, so that frequently two or more answers could conceivably be correct. A third problem was that some directions were incomplete, omitting information on how and where the student should mark his response. It was suggested that directions needed more editorial work and that instructions to the student should be appropriate to the listening ability of that grade.

New tests will be developed for future statewide assessments, so the problems encountered during this initial attempt at statewide assessment can be used to avoid the same problems with future tests. For example, contracts for test construction now contain more criteria for clarity and understandability of illustrations and directions.

INSTRUCTIONS AND GENERAL INFORMATION

While the instructions for all phases of assessment were adequate to get the job done, they did not cover certain situations, which created unforeseen difficulties. For example, some schools have two numbers, one assigned by the district and the other by the Department of Education. Materials were packed and shipped using the Department of Education school number, and it was anticipated that these numbers would be used by school personnel when filling out the back of the test form. However, some schools wrote in the district school number on the test forms. Since the test scorers used the school number to identify the school and the tests, they lost almost three weeks trying to match tests with schools.

The problem was compounded by the fact that some schools tested more students than they were asked to test, an action not specifically forbidden in the instructions. This resulted in additional verification to determine whether the schools had overtested or whether school enrollment had increased.

These and similar minor problems suggest a need for greater attention to preparation of all instructions. A specific recommendation for improving the instructions was to compile the responsibilities of all individuals involved in assessment into one document. The document would contain a flowchart and other devices for clarifying who is to do what, when. District coordinators also identified a need for more information, such as inservice training materials, articles and news releases, to help them explain assessment to school personnel and to the public.

PACKAGING AND DISTRIBUTION OF TEST MATERIALS

One area in which the procedures were less successful was in the packaging and distribution of test materials. The procedures caused no major delays or missing data, but they were not maximally efficient, causing extra effort for the district coordinators. Because a temporary, inexperienced staff was packaging the materials under extreme time pressure, the materials shipped to the schools did not always agree with the packing slips, which necessitated additional checking at the district and school level. The instructions to the school coordinator did not specify that the boxes should be saved and used to return the materials to the Department of Education, so some districts had difficulty finding boxes of the right sizes to repackage the material. Finally, coordinators encountered difficulties in finding a reasonable method of returning materials, since in some cases they were not allowed to use the educational rate because the tests had been written in.

The problems of this initial year of testing have been carefully examined by the Department of Education. Because of the experience gained this year, more effective procedures will be utilized in the 1972-73 assessment and most of the problems encountered this year should be eliminated.

APPENDIX A

THE EDUCATIONAL ACCOUNTABILITY ACT OF 1971

Chapter 229.57, Florida Statutes

Section 1. Short Title.--This act shall be known and may be cited as "The Educational Accountability Act of 1971."

Section 2. Purposes; Intent.--The purposes of this act are to provide for the implementation and further development of educational assessment procedures as required by section 9(1), chapter 70-399, Laws of Florida, and the plan for educational assessment in Florida, developed by the commissioner of education pursuant to this chapter;

(a) To provide for the establishment of educational accountability in the public educational system of Florida;

(b) To assure that education programs operated in the public schools of Florida lead to the attainment of established objectives for education:

(c) To provide information for accurate analysis of the costs associated with public education programs: and

(d) To provide information for an analysis of the differential effectiveness of instructional programs.

Section 3. Educational Accountability Program.--The commissioner of education is directed to implement a program of educational accountability for the operation and management of the public schools, which shall include the following:

(a) Pursuant to section 229.053(2)(e), Florida Statutes, the commissioner, with the approval of the state board of education, shall, no later than November 1, 1972, and each year thereafter, establish basic, specific, uniform statewide educational objectives for each grade level and subject area, including, but not limited to, reading, writing, and mathematics, in the public schools.

(b) The commissioner shall develop and administer a uniform, statewide system of assessment based in part on criterion-referenced tests and in part on norm-referenced tests to determine periodically pupil status, pupil progress, and the degree of achievement of established educational objectives.

(c) The commissioner shall make an annual public report of the aforementioned assessment results. Such report shall include, but not be limited to, a report of the assessment results by grade and subject area for each school district and the state, with an analysis and recommendation concerning the costs and differential effectiveness of instructional programs.

(d) The school board of each district shall by the 1973-1974 school year make an annual public report of the aforementioned assessment results which shall include pupil assessment by grade and subject area for each school in the district. A copy of the district's public report shall be filed with the commissioner of education.

(e) The commissioner, with approval of the state board of education, shall by the 1973-1974 school year, develop accreditation standards based upon the attainment of the established educational objectives.

Section 4. Implementation.--This act shall apply to the subject area of reading by the 1971-1972 school year and the subject areas of writing and mathematics by the 1972-1973 school year and shall include other subject areas by the 1973-1974 school year.